

# A comparison of coronal mass ejections identified by manual and automatic methods

S. Yashiro<sup>1,2</sup>, G. Michalek<sup>1,2,3</sup>, and N. Gopalswamy<sup>2</sup>

<sup>1</sup>Catholic University of America, Washington, D.C., 20064, USA

<sup>2</sup>NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA

<sup>3</sup>Astronomical Observatory of Jagiellonian University, Krakow, Poland

Received: 9 October 2007 – Revised: 12 March 2008 – Accepted: 18 July 2008 – Published: 15 October 2008

**Abstract.** Coronal mass ejections (CMEs) are related to many phenomena (e.g. flares, solar energetic particles, geomagnetic storms), thus compiling of event catalogs is important for a global understanding these phenomena. CMEs have been identified manually for a long time, but in the SOHO era, automatic identification methods are being developed. In order to clarify the advantage and disadvantage of the manual and automatic CME catalogs, we examined the distributions of CME properties listed in the CDAW (manual) and CACTus (automatic) catalogs. Both catalogs have a good agreement on the wide CMEs (width $>120^\circ$ ) in their properties, while there is a significant discrepancy on the narrow CMEs (width $\leq 30^\circ$ ): CACTus has a larger number of narrow CMEs than CDAW. We carried out an event-by-event examination of a sample of events and found that the CDAW catalog have missed many narrow CMEs during the solar maximum. Another significant discrepancy was found on the fast CMEs (speed $>1000$  km/s): the majority of the fast CDAW CMEs are wide and originate from low latitudes, while the fast CACTus CMEs are narrow and originate from all latitudes. Event-by-event examination of a sample of events suggests that CACTus has a problem on the detection of the fast CMEs.

**Keywords.** Solar physics, astrophysics, and astronomy (Flares and mass ejections)

## 1 Introduction

Since the discovery of the coronal mass ejection (CME) phenomenon in 1971 (Tousey, 1973), CMEs have been observed by several space-borne instruments, i.e. the seventh Orbiting Solar Observatory (OSO-7) coronagraph, the Apollo Telescope Mount (ATM) coronagraph on board Skylab, the Sol-

wind coronagraph on board the P78-1 satellite, the Coronagraph/Polarimeter (CP) on board Solar Maximum Mission (SMM), the Large Angle and Spectrometric Coronagraph (LASCO) on board the Solar and Heliospheric Observatory (SOHO), and Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) on board Solar TERrestrial Relations Observatory (STEREO). The ground-based instruments, e.g. the Mauna Loa K-Coronameter, have also observed CMEs. For more than three decades, the identification of the CMEs has been carried out by human eyes and the event catalogs have been compiled manually (e.g. Howard et al., 1985; St. Cyr and Burkepile, 1990; Burkepile and St. Cyr, 1993; Yashiro et al., 2004; Gopalswamy et al., 2008).

LASCO operators have checked observed images daily to watch the instrument status and to broadcast alerts of halo CMEs for space weather purposes. They have also been compiling preliminary CME lists, which are available at the LASCO homepage (<http://lasco-www.nrl.navy.mil/>). St. Cyr et al. (2000) carried out the basic measurements, e.g. CME speed and angular width, and described the basic properties of the LASCO CMEs covering the first three years of the LASCO operation. Based on the preliminary lists, we have been measuring basic CME properties (onset time, speed, angular width, and so on), and making them available the Internet at the CDAW Data Center ([http://cdaw.gsfc.nasa.gov/CME\\_list/](http://cdaw.gsfc.nasa.gov/CME_list/)). In addition to the measurements (in digital format as well as plots), the catalog also contains various types of movies made from the coronagraph, so users can directly check the reality of the CMEs and the listed measurements. In the paper we refer to this catalog as the CDAW catalog. Detailed description of the CDAW catalog can be found in Yashiro et al. (2004) and Gopalswamy et al. (2008).

Thanks to the extended operation of the SOHO mission and the high capability of the LASCO coronagraph, more than eleven thousand CMEs have been recorded from January 1996 to December 2006. In addition to the manual detection and measurement of CMEs, there have also been

---

Correspondence to: S. Yashiro  
([seiji.yashiro@nasa.gov](mailto:seiji.yashiro@nasa.gov))

several attempts to automatically catalog CMEs from the LASCO data. One such program is the “Computer Aided CME Tracking”, or CACTus (Berghmans, 2002; Berghmans et al. 2002). Robbrecht and Berghmans (2004) and Robbrecht et al. (2006) improved CACTus and a catalog of CMEs identified by CACTus is available on line at <http://www.sidc.be/cactus/>. Boursier et al. (2005, 2006) have developed another method known as the “Automatic Recognition of Transient Events and Marseille Inventory from Synoptic maps” (ARTEMIS) to identify CMEs from synoptic maps. A catalog of CMEs identified from ARTEMIS will be soon available at <http://lascor.oamp.fr/lasco/>. Qu et al. (2006) presented a method to detect, characterize and classify the CMEs, and have published their results at <http://filament.njit.edu/detection/vso.html>. Another method is the “Solar Eruptive Event Detection System” (SEEDS) developed by Olmedo et al. (2008). Their website is at <http://spaceweather.gmu.edu/seed/>.

The manual CME identifications depend on observers. Even for the same observer, the identification ability is not constant. Thus manual catalogs are “subjective”. Generally speaking, all automatic methods identify more CMEs than manual detection. However there is no one-to-one correspondence among automated catalogs. In this context, the automated catalogs are also somewhat subjective because the CME identifications depend on the computer programs. One advantage of the automated methods is that everyone can obtain the same results using the same procedure (in this sense the automated methods are objective). This advantage is valid only when an automated method can identify CMEs and compile the CME characteristics properly. However, the validation is important for both manual and automatic methods. The CDAW catalog (manual) has been extensively used by the scientific community and the CACTus catalog (automatic) has also been well developed. In this paper we compare these two catalogs to clarify their advantages and disadvantages.

## 2 CME identifications and properties

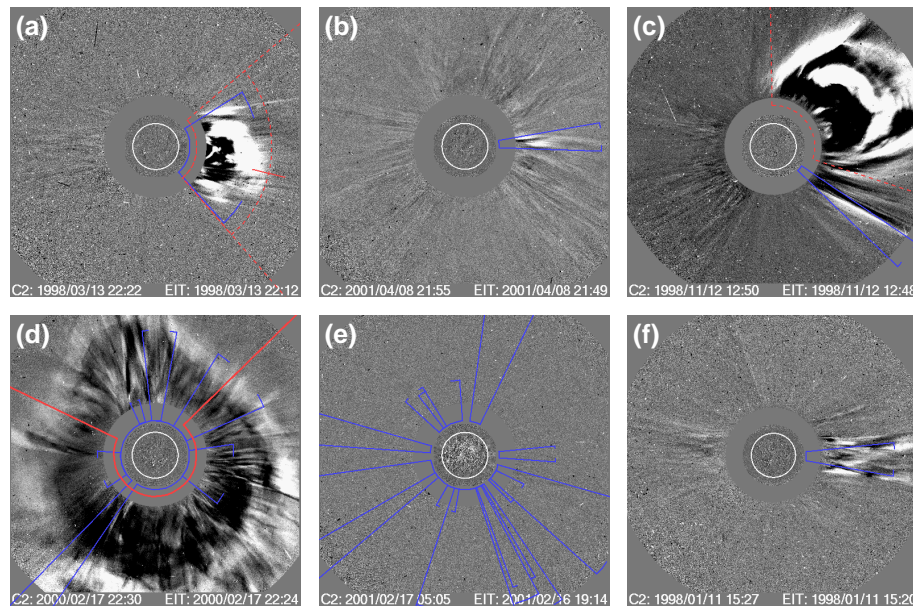
CMEs are eruptions of plasma from the Sun at speeds from 10–3500 km/s (Gopalswamy, 2004; Yashiro et al., 2004). Observationally the CME phenomenon is defined as a new, discrete, bright white-light feature appearing in the corona on time scales of minutes to hours and moving outward (Hundhausen, 1993; St. Cyr et al., 2000; Yashiro et al., 2004). This has been used for a long time, but some difficult cases arise when we identify CMEs using this definition. One such case is to decide whether an erupting feature preceded by a large CME is a distinct event, or a part of the preceding CME. For the CDAW catalog, we have listed a feature as a CME only if the feature has clear signatures such as the CME three-part structure (Illing and Hundhausen, 1985); the feature erupts faster than the preceding CME; the feature and the preced-

ing CME are associated with the different X-ray flares. If the feature does not have any of these signatures, we do not list it in the CDAW catalog though it might be a separate CME.

As of August 2007 the CDAW catalog has monthly event lists from January 1996 to December 2006 while the CACTus catalog has from April 1997 to March 2007. Therefore, we compare the CMEs in the overlapping period (April 1997–December 2006). Approximately one-fourth of the events in the CACTus catalog are labeled by “marginal case” for the events whose signal is near the noise level (see <http://sidc.oma.be/cactus/scan/>). We eliminated these events from our analysis. In the study period, the CDAW Catalog has 11 394 CMEs and the CACTus catalog has 17 341 CMEs (excluding the marginal cases).

Figure 1a shows a CME listed in the CDAW (red) and CACTus (blue) catalogs. The side edges were computed from the central position angle (CPA) and angular width. For the CDAW CME, the outer arc indicates the height of the leading edge obtained from height-time digital data. As the CACTus catalog does not have the CME height data, we estimated the heights from the CME onset ( $t_0$  in CACTus) and speed ( $v$  in CACTus). The CME height when it first appears in the LASCO C2 field of view was assumed to be 2.5 solar radii. This is a rough approximation, thus the blue arc for a CACTus CME might sometimes lie above the CME leading edge (see Fig. 1b). We created monthly movies in quick-time format, which are available at CDAW website (<http://cdaw.gsfc.nasa.gov/publications/soho20movie/>).

CACTus detects more CMEs than CDAW lists during the same interval, meaning that either CDAW missed some true CMEs, CACTus has false CME detections, or both catalogs have errors. The former case is shown in Fig. 1b: a CME listed in the CACTus catalog but not in the CDAW catalog. This is due to human error, which usually can happen for minor CMEs. The latter cases are shown in Fig. 1c–e. In this paper we refer to a CACTus event as a false event when it is not recognized as a CME after careful inspection by human eyes. CACTus detects streamer deflections pushed by a large CME as separate events (Fig. 1c). However, streamer deflections are not a “mass ejection” as no mass is ejected from the Sun, thus we refer to the event as a false CME. CACTus occasionally detects internal parts of a CME as separate events. For example, CACTus detected the halo event in Fig. 1d as a halo CME at 19:54 UT, and then detected 6 following events during 21:54 UT–23:06 UT. We think CACTus detected the motion of the internal structures of the halo. These are not separate CMEs hence are labeled as false. For some CACTus CMEs we could not recognize any features moving outward even after we watched the LASCO movies carefully. CACTus might have detected background solar wind flows. Some of them are doubtful in their reality because many events are detected during a short period (Fig. 1e). CACTus occasionally detects gusty flows in the aftermath of a preceding large CME (Fig. 1f). These could be a part of the preceding CMEs or separate CMEs. In this case we can not determine whether



**Fig. 1.** CMEs in the CDAW (red) and CACTus (blue) catalogs. (a) A CME in both catalogs. The listed properties for the CME are almost the same in the two catalogs. (b) A CME listed in CACTus but not in CDAW. (c–e) False CACTus CMEs. (f) Gusty flows in the aftermath of a large CME. This is a possibly true CACTus CMEs. The monthly movies for sampled period are available at CDAW website (<http://cdaw.gsfc.nasa.gov/publications/soho20movie/>).

the event is true or false, thus we refer to them as possibly true CMEs.

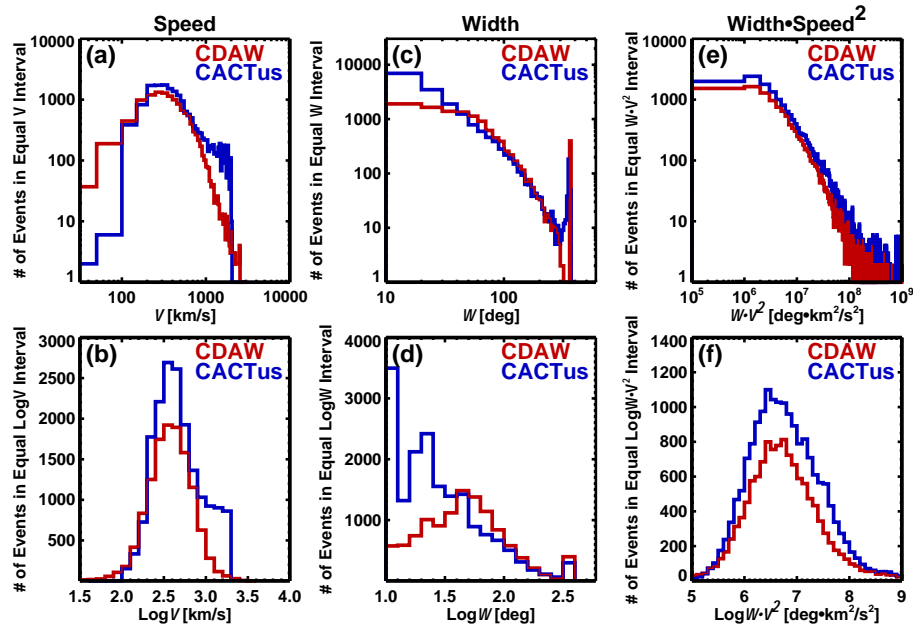
Basic properties of CMEs are listed in both catalogs, but some parameters are not. For example, the CDAW catalog lists acceleration for each CME but CACTus does not. In the CACTus catalog, each CME has multiple speeds for each position angle within the CME span.

Speed is one of the most important parameters characterizing the CME. For the CDAW CMEs, we used linear speeds obtained by fitting a straight line to the height-time measurements. The CME height is the distance from the Sun center to the leading edge (LE) of the CME, thus speeds in the CDAW catalog represent motion of the CME LE at its fastest section. On the other hand, CACTus detects a CME as the bright line in height-time slice images obtained using the Hough transform technique, thus the CACTus speeds represent motion of the brightest structure of CMEs. Therefore, for the same CME, the CACTus speed is usually lower than the CDAW derived speed and this trend was confirmed by Robbrecht and Berghmans (2004). CACTus measures speeds at different position angles within the CME span since the height-time slice images are composed along multiple position angles. The CACTus catalog lists the minimum, maximum, and median speeds. On the basis of the idea that similar parameters should be used for the comparison, we should select the maximum speed in CACTus since height-time measurements in CDAW are made at the PA where the CME's leading edge moves the fastest. However, we used the median speeds for this analysis because CACTus uses them as the representative speed of CMEs.

### 3 Speed and width

Before describing our results we should emphasize here that the distributions shown in the paper are the distributions of “apparent” CME parameters because measurements are made in the sky plane. We did not correct for projection effects. In addition, due to the nature of CME observations, not all weak CMEs are detectable by LASCO (Yashiro et al., 2005). The observed parameters are biased. These projection effects and biases equally affect both catalogs, hence are not important for the comparison, but the true distributions would be different.

Figure 2 shows distributions of speed ( $V$ ), angular width ( $W$ ), and the product  $W \times V^2$  of CMEs listed in both catalogs. We use  $W \times V^2$  as a proxy to the kinetic energy since the width is proportional to the CME mass (Gopalswamy, 2006b). Red and blue lines correspond to the distributions of CDAW and CACTus CMEs, respectively. The same color code is used in all other figures. The distributions in the top row are shown in the log-log plots, but the numbers of events are counted by equal intervals on a linear scale. Therefore a straight line in the plots suggests a power-law distribution. In the bottom row, the logarithms of the parameters are used to count the numbers of events, i.e. bin sizes are equal on a log scale. The plots are used to investigate whether a parameter obeys the log-normal distribution. For example, the speed distribution of CDAW CMEs (the red line in Fig. 2b) is Gaussian, indicating that the CDAW CME speed distribution is log-normal (Aoki et al., 2003; Yurchyshyn et al., 2005).



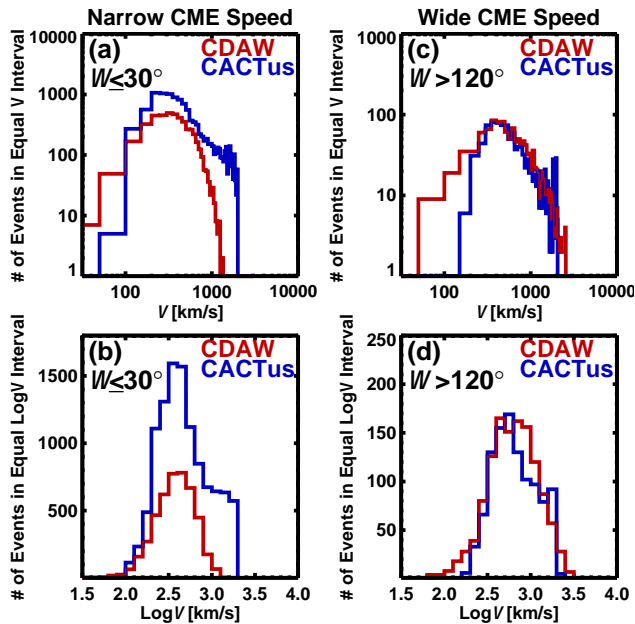
**Fig. 2.** Distributions of Speed ( $V$ ; left), width ( $W$ ; center), and  $W \times V^2$  (right) of CDAW (red) and CACTus (blue) CMEs. The parameter  $W \times V^2$  is used as a proxy of the kinetic energy. The number of events in each bin is counted by equal intervals on a linear scale (top) and on a log scale (bottom), which is usually used to investigate the power-law and log-normal distributions, respectively.

The speed distributions of CDAW and CACTus CMEs are shown in Fig. 2a and b. Both the distributions are peaked around 300 km/s, but the shapes are different. Although the speed distribution of CDAW CMEs obeys the log-normal distribution, that of CACTus CMEs does not. The difference might arise because the two catalogs list speeds corresponding to different parts of the CMEs. The CDAW catalog has more slow CMEs ( $V < 100$  km/s) and the CACTus catalog has more fast CMEs ( $V > 1000$  km/s). The lack of slow CMEs in CACTus is due to a technical restriction: CACTus rejects any moving features appearing continuously in more than 250 LASCO images. This corresponds to CMEs with speed less than 100 km/s during the normal LASCO cadence (three C2 and two C3 images per hour). Therefore slow CMEs ( $V < 100$  km/s) are only occasionally detected when LASCO images have poor cadence or has data gaps (Robbrecht and Berghmans, 2004). CACTus has 2655 fast CMEs occurring from April 1997 to December 2006 while CDAW has only 494 fast ones during the same period. The extreme excess of fast CMEs in CACTus is unexpected because (1) for a given CME the CACTus speed is expected to be lower than the CDAW speed, and (2) fast CMEs are easy to identify by human eyes hence are not missed in the CDAW catalog. This difference is very important for space weather applications because fast CMEs drive shocks, which accelerate particles. In order to see the difference in detail, we have shown the speed distributions of narrow ( $W \leq 30^\circ$ ; left) and wide ( $W > 120^\circ$ ; right) CMEs in Fig. 3. In the same way as Fig. 2, the numbers in each bin are counted by equal

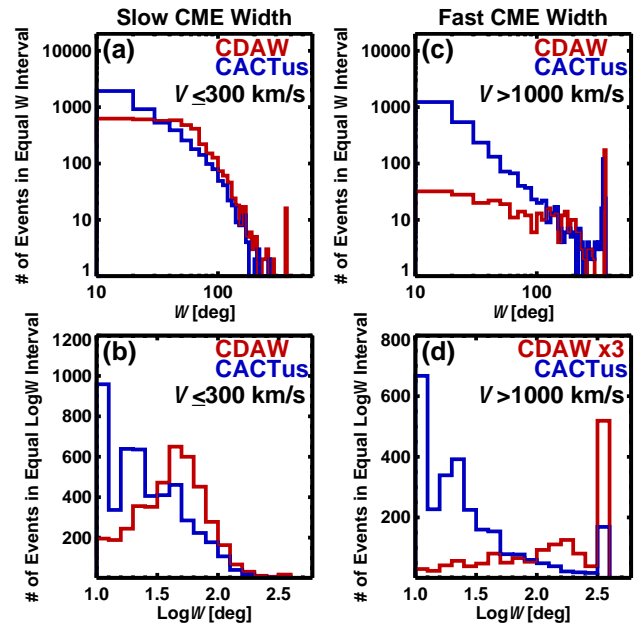
intervals on a linear scale (top) and on a log scale (bottom). In Fig. 3c and d, we see that the speed distributions of the wide CMEs listed in CDAW and CACTus are almost identical except for the slower CMEs: CDAW has more slow and wide CMEs. Two major differences between the catalogs are in the number of the narrow CMEs in two speed ranges, 200–500 km/s and  $> 1000$  km/s: CACTus has 3218 and 1755 additional CMEs, respectively.

The width distributions are shown in Fig. 2c and d. Both the distributions are almost the same except for the number of narrower CMEs. The difference is prominent in the bins below 30 degrees. CDAW has 3982 narrow CMEs while CACTus has more than twice that number (10 796). Except for the amplitude of the increase, this is expected because human eyes are likely to miss weak CMEs. Thanks to the increase of narrow CMEs, the width distribution of CACTus CMEs obeys a power-law distribution. Figure 4 shows the angular width distributions of slow ( $V \leq 300$  km/s; left) and fast ( $V > 1000$  km/s; right) CMEs. For the slow CMEs, CDAW has more wide CMEs while CACTus has more narrow CMEs. The difference is more prominent for the fast CMEs shown in Fig. 4d. The majority of fast CMEs in the CDAW catalog are halo CMEs, while those in CACTus are narrow CMEs.

Figure 2e and f shows the distributions of  $W \times V^2$ , as a proxy to the CME kinetic energy. Interestingly, both distributions are very similar and obey a power-law even though there are significant differences in the speed and width distributions. We obtained a power-law index  $\alpha = 1.69$



**Fig. 3.** Speed Distributions of narrow ( $W < 30^\circ$ ; left) and wide ( $W > 120^\circ$ ) CMEs. In the same way as Fig. 1, the number of events in each bin is counted by equal intervals on a linear scale (top) and on a log scale (bottom).



**Fig. 4.** Angular width distributions of slow ( $V < 300$  km/s; left) and fast ( $V > 1000$  km/s; right) CMEs. Red and Blue lines are for CDAW and CACTus events, respectively. In panel d, the numbers of CDAW events are multiplied by 3 for clarity.

**Table 1.** Narrow CACTus CMEs<sup>a</sup>.

Speed Range in km/s	Number	True	Non-Narrow	Possible	False
0–300	2935	38%	12%	20%	30%
300–500	3399	53%	7%	9%	31%
500–1000	2636	59%	3%	5%	33%
>1000	1826	27% <sup>b</sup>	6%	13%	54%
ALL	10796	46%	7%	12%	35%

<sup>a</sup> The percentages were estimated from 100 sample events in each speed range.

<sup>b</sup> The success rate would decrease to 3% if non-fast CMEs were excluded.

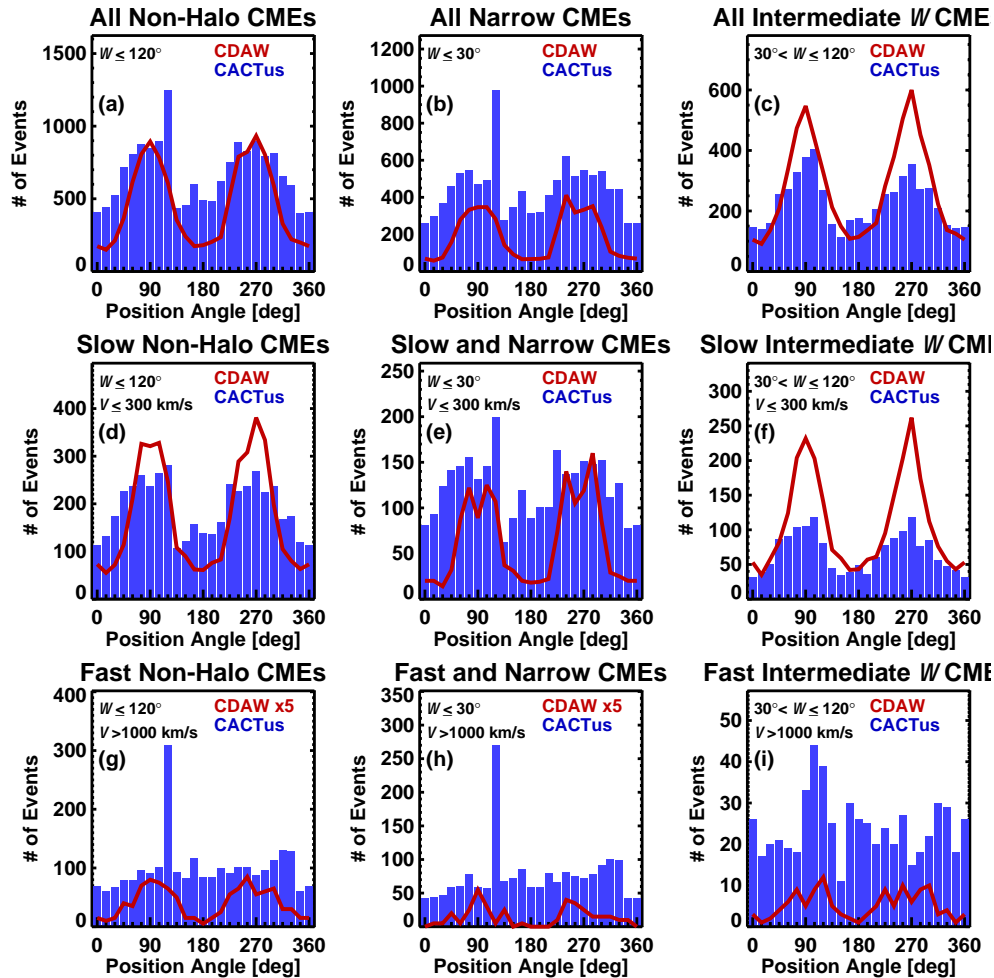
for the CDAW and  $\alpha=1.71$  for the CACTus distributions. Yashiro et al. (2006) found that the frequency distributions of flares with and without CMEs as a function of the flare fluence (total X-ray flux) were power laws with different indices:  $\alpha=1.79 \pm 0.05$  for the CME-associated flares and  $\alpha=2.47 \pm 0.11$  for the CME-less flares. The former is somewhat similar to the power-law index of the  $W \times V^2$  distribution. However, as we described in the beginning of this section, many weak CMEs originating away from the limb are not observed. Therefore, the true  $W \times V^2$  distribution might be steeper than observed.

### 3.1 Event-by-event examination of a sample of events

We found a significant discrepancy in the number of narrow CMEs, so we carried out an event-by-event examination of a sample of events to find the reasons. In order to check differences between fast and slow ones, we divided the events into four speed ranges: 0–300 km/s, 301–500 km/s, 501–1000 km/s, and >1000 km/s. From the 2935 narrow CACTus CMEs in the speed range 0–300 km/s, we picked every 29th event (in occurrence time order), and obtained 100 samples. In the same manner we picked 300 samples from other three speed ranges. Out of the 10 796 narrow CMEs listed in the CACTus catalog, 400 samples were visually examined and classified into four groups: (1) true CMEs, (2) non-narrow true CMEs, (3) possible CMEs, and (4) false CMEs. The “non-narrow CMEs” are those whose detection as CMEs is no problem but their measured width is lower than half the width of the corresponding CDAW CMEs. The “possible CMEs” are CACTus events whose reality as CMEs is difficult to determine.

The results of the visual examination are summarized in Table 1. In the high speed range (>1000 km/s), 54 out of 100 sample CMEs were found to be false. We found several internal parts of a single fast and wide CME were falsely detected as many narrow events (see Fig. 1d). We confirmed 27 CACTus events were true and their width measurement had no problem. However we found 24 out of them had a problem on the speed detection: Their CACTus speed was





**Fig. 5.** Central position angle distributions of (a) all non-halo CMEs, (b) fast non-halo CMEs, (c) narrow CMEs, and (d) fast and narrow CMEs, which are listed in CACTus (blue bars) and CDAW (red lines) catalogs. The numbers of fast CDAW CMEs (panels g and h) are multiplied by 5. The anomaly peaks in CACTus are shown in 120° bin, which correspond to the coronagraph pylon of LASCO.

higher than 1.5 times of CDAW derived speed. The remaining 3 were truly fast and narrow CMEs. Therefore, we can estimate that only 55 out of the 1826 fast and narrow CACTus CMEs are real. The number of fast and narrow CMEs in CDAW is close to 70. Therefore, the speed distributions of the fast CACTus CMEs shown in Fig. 4c and d are incorrect. There is no evidence contradicting the result that the majority of the fast CMEs are wide.

The fractions of true, non-narrow, possible, and false CMEs for all narrow CACTus events in the last row are the weighted average of the fractions in the four speed ranges. We estimated 776 (or 7%) narrow CACTus CMEs are not narrow in CDAW and 3779 (or 35%) narrow CACTus CMEs are false. These are not negligible numbers and hence heavily influence the shape of the width distribution for CACTus CMEs. We estimated that 4966 (or 46%) of narrow CACTus CMEs are true, and 1295 (or 12%) are possibly true, indicating that there are 1000–2300 CMEs more than in the CDAW

catalog (3982). Therefore, a significant number of narrow CMEs are not listed in the CDAW catalog.

#### 4 Location

The location of a CME is represented by the central position angle (CPA), which is defined as the mid-angle of the two side edges of the CME in the sky plane. Position angle (PA) is measured counterclockwise from Solar North in degrees. The PAs 0°, 90°, 180°, 270°, and 360° correspond to the north pole, eastern equator, south pole, western equator, and north pole, respectively. Figure 5 shows the CPA distributions of CDAW CMEs (red lines) and CACTus CMEs (blue bars). The partial and full halos ( $W > 120^\circ$ ) are excluded from this analysis. For these CMEs it is likely that CPAs are highly projected, thus the CPAs do not indicate the CME location properly.

The CPA distributions of the non-halo CMEs are shown in Fig. 5a. The CACTus CMEs show an anomalous peak in the  $120^\circ$  bin, which corresponds to the coronagraph pylon of LASCO. It is mentioned in the CACTus homepage that there are “noisy events” around the pylon due to both a lower signal and smaller signal-to-noise ratio. Except for this, both the distributions have two peaks in the equatorial region (corresponding to  $90^\circ$  and  $270^\circ$  CPAs), and their amplitudes are nearly the same. The differences between CDAW and CACTus CMEs arise in the number of CMEs at mid- and high-latitudes. We divided the events into two groups, narrow CMEs ( $W \leq 30^\circ$ ; Fig. 5b) and intermediate-width CMEs ( $30^\circ < W \leq 120^\circ$ ; Fig. 5c). The anomalous peak at  $120^\circ$  is prominent in Fig. 5b, but not seen in Fig. 5c, suggesting that the false CACTus CMEs corresponding to the coronagraph pylon region are narrow.

Figure 5d–f shows the CPA distributions of the slow CMEs ( $V \leq 300$  km/s) in the CDAW and CACTus catalogs. All the distributions have two peaks near the equator, but the CACTus CMEs are widely distributed to high latitudes compared to the CDAW CMEs. From Fig. 5f we see that CDAW has more slow and intermediate-width CMEs ( $V \leq 300$  km/s;  $30^\circ < W \leq 120^\circ$ ).

The CPA distributions of fast CMEs ( $V > 1000$  km/s) in the CDAW and CACTus catalogs are shown in Fig. 5g. In order to see their differences clearly, the number of the CDAW events are multiplied by an appropriate factor, which is denoted in each panel. The anomalous peak at  $120^\circ$  is prominent, suggesting that the false CACTus CMEs corresponding to the coronagraph pylon are fast. The distribution of CDAW CMEs has two peaks at the eastern and western parts of the equator while the CACTus distribution is relatively flat. The CPA distributions of the fast and narrow CMEs ( $V > 1000$  km/s;  $W \leq 30^\circ$ ) and fast and regular width CMEs ( $V > 1000$  km/s;  $30^\circ < W \leq 120^\circ$ ) are shown in Fig. 5h and i, respectively. Again, the CDAW CMEs are mainly distributed around the equator, but the number of events is very small compared to that of the CACTus CMEs. The fast CMEs in CACTus are distributed at all latitudes and there is no preferable position. This difference is important for the origin of the fast CMEs. The CDAW result suggests that the fast CMEs mainly originate from low latitudes, i.e. from the active region belts, while the CACTus result suggests that a significant number of fast CMEs also originate from high latitudes.

#### 4.1 Event-by-event examination of a sample of events

In the location distribution of fast CMEs there is a significant discrepancy between the two catalogs: fast CDAW CMEs mainly originate from low latitudes while fast CACTus CMEs originate from all latitudes. We carried out an event-by-event visual examination of the fast CACTus CMEs originating from low ( $0\text{--}30^\circ$ ), middle ( $31\text{--}60^\circ$ ), and high ( $61\text{--}90^\circ$ ) latitudes. We excluded the CACTus events in the  $120^\circ$

**Table 2.** Fast CACTus CMEs<sup>a</sup>.

Apparent Latitude	Number	True	Non-Fast	Possible	False
0–30	714	6%	36%	14%	44%
31–60	697	10%	24%	14%	52%
61–90	672	1%	19%	11%	69%
ALL	2083	6%	26%	13%	55%

<sup>a</sup> The percentages were estimated from 100 sample events in each latitude range.

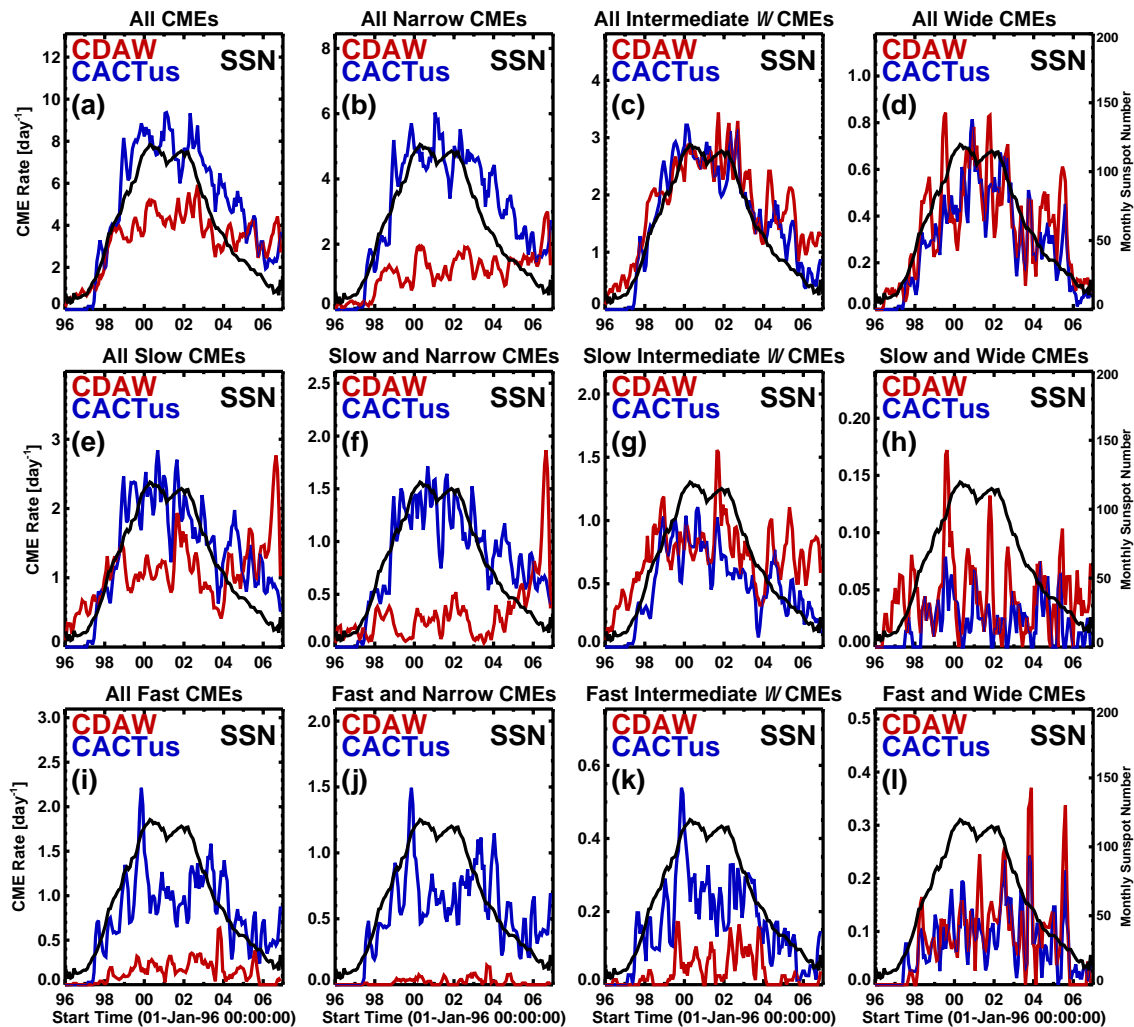
degree bin from the examination because their reality as CMEs is already doubtful. Again 100 events are sampled from each range. After checking the LASCO movies, we classified the 300 fast CACTus events into four groups: (1) true CMEs, (2) non-fast true CMEs, (3) possible CMEs, and (4) false CMEs. The second group “non-fast true CMEs” is for a CACTus event whose detection as a CME is no problem but the measured speed is higher than 1.5 times of the CDAW derived speed.

The results of the visual examination are shown in Table 2. We find that 6% of the fast CACTus CMEs are real CMEs, and 26% are non-fast real CMEs, and 13% are non-fast possible CMEs. We identified the remaining 55% are false detection by CACTus. The detection is worse for the high-latitude fast events: 1% are real and 69% are false. The CACTus catalog has a large number of high-latitude fast CMEs because of problems either in the speed measurements or in the identification of fast CMEs.

## 5 Solar cycle variation

The occurrence rates of the CMEs listed in the CDAW and CACTus catalogs are shown in Fig. 6. Again, the red and blue lines correspond to the CDAW and CACTus catalogs, respectively. For reference, the monthly smoothed sunspot number compiled by the Solar Influences Data analysis Center (SIDC; <http://sidc.oma.be>), Belgium is plotted (black). The occurrence rate was determined by dividing the number of CMEs in each Carrington rotation by total observation time in days. The obtained rate was smoothed over three Carrington rotations to reduce the short-term fluctuations. Similar to Fig. 5, we divided the CMEs by their speed and width. The second, third and fourth columns are for narrow ( $W \leq 30^\circ$ ), intermediate-width ( $30^\circ < W < 120^\circ$ ), and wide ( $W > 120^\circ$ ) CMEs, respectively, and the middle and bottom lines for slow ( $V \leq 300$  km/s) and fast ( $V > 1000$  km/s) CMEs, respectively.

The occurrence rates of all CMEs listed in CDAW and CACTus are shown in Fig. 6a. The rate of CACTus CMEs closely follows the sunspot number, but the CDAW rate does not. The discrepancy is much clearer in the rate of narrow



**Fig. 6.** Daily CME rate as a function of time. Red and blue lines are for the CDAW and CACTus catalogs, respectively (left scale). Black lines are smoothed monthly sunspot number (right scale) for the reference of the solar cycle.

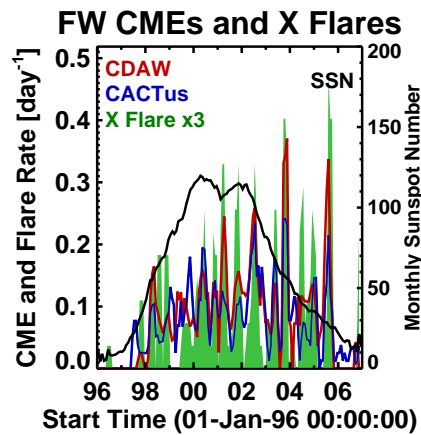
CMEs (Fig. 6b) and slow, narrow CMEs (Fig. 6f). The number of narrow CMEs listed in the CDAW catalog keeps increasing even during the declining phase. The CDAW observers have labeled the quality of the “tracking feature” of a CME as belonging to one of the following six categories: excellent, good, typical, fair, poor, and unable to measure (Yashiro et al., 2004). During 2000–2002, 50% of the CDAW CMEs are labeled as “poor” or “very poor” while the rate is 90% for 2006. It is reasonable to think that the CDAW catalog might have missed many faint CMEs (poor, very poor) during the solar maximum. In Fig. 6d we have shown the occurrence rates of wide CMEs ( $W > 120^\circ$ ) listed in the CDAW and CACTus catalogs. Their solar cycle variations are very similar over the whole of the cycle. Overall it appears that both the rates follow the sunspot number, but there are some excesses during the declining phase (2003–2006). The excesses were caused by the existence of the super active re-

gions (Gopalswamy et al., 2006). Figure 6l shows the occurrence rate of fast and wide CMEs ( $V > 1000$  km/s;  $W > 120^\circ$ ). The rates from CDAW and CACTus are very similar and do not follow the sunspot number. Figure 7 is a larger version of Fig. 6l with occurrence rate of X-class flares (green shade) for additional reference. Again, the rate of X-class flares does not follow the sunspot number, and many peaks are well correlated with that of fast and wide CMEs. Any correlation of X-class flares with the smoothed sunspot number is overwhelmed by two intense periods of X-class flare activity in the declining phase of the solar cycle.

## 6 Summary and conclusions

We compared the statistical properties of CMEs listed in the CDAW (manual) and CACTus (automatic) catalogs. Both catalogs have a good agreement on wide CMEs ( $W > 120^\circ$ )





**Fig. 7.** Daily rate of fast wide CMEs and X-class flares. Red and blue lines are for the CDAW and CACTus catalogs, respectively, and Green shade is for the X-class flares (left scale). Black lines are smoothed monthly sunspot number (right scale) for the reference of the solar cycle.

in their properties: the speed distribution (Fig. 3c and d) and the solar cycle variations of the occurrence rate (Fig. 6d). The major discrepancies between the two catalogs are:

1. Identification of the narrow CMEs ( $W \leq 30^\circ$ ; Fig. 2c). CACTus has 10 796 narrow CMEs from April 1997 to December 2006, while CDAW has only 3982 during the same period. By event-by-event visual examinations of a sample of events, we estimated that the CDAW catalog might have missed 1000–2300 narrow CMEs, while the CACTus catalog lists 3800 false events. In terms of the solar cycle variation of the CME rate (Fig. 6b), CACTus follows the sunspot number but CDAW does not, indicating that CDAW might have missed some CMEs during the maximum phase. However, the CME rate need not closely follow the sunspot cycle, because CMEs also originate from non-sunspot regions especially during solar maximum from polar crown filaments (Gopalswamy et al., 2003a, b).
2. Identification of fast CMEs ( $V > 1000$  km/s; Fig. 2b). During the study period, CDAW has 494 fast CMEs, while CACTus has 2655. Discrepancies are also found in the width and location distributions: Most of the fast CMEs in the CDAW catalog have large angular widths while those in CACTus are narrow (Fig. 4d). The fast CDAW CMEs are mainly located at low latitudes, i.e. the active region belt, while the fast CACTus CMEs originate from all latitudes (Fig. 5c). By the event-by-event visual examinations of a sample of events, we found that only 6% of the fast CACTus CMEs are truly fast: therefore the distributions of fast CACTus CMEs are unreliable.

Obviously both CDAW and CACTus catalogs have some problems. The CDAW catalog must have missed some narrow CMEs during the period before year 2004 as evidenced by the increase in 2005.

**Acknowledgements.** We acknowledge the SIDC team for providing the International Sunspot Number, which is available at <http://www.sidc.be/sunspot-data/>. SOHO is a project of international cooperation between ESA and NASA. The LASCO data used here are produced by a consortium of the Naval Research Laboratory (US), Max-Planck-Institut für Aeronomie (Germany), Laboratoire d'Astronomie (France), and the University of Birmingham (UK). Part of this effort was supported by NASA (NNG05GR03G).

Topical Editor R. Forsyth thanks two anonymous referees for their help in evaluating this paper.

## References

- Aoki, S., Yashiro, S., and Shibata, K.: The Log-Normal Distributions of Coronal Mass Ejection-Related Solar Flares and the Flare/CME Model of Gamma-Ray Bursts, in: *Proceedings of the 28th International Cosmic Ray Conference*, 2729–2732, 2003.
- Berghmans, D.: Automated detection of CMEs, ESA SP-506, 85–89, 2002.
- Berghmans, D., Foing, B. H., and Fleck, B.: Automated detection of CMEs in LASCO data, ESA SP-508, 437–440, 2002.
- Boursier, Y., Llebaria, A., Goudail, F., Lamy, P., and Robelus, S.: Automatic detection of coronal mass ejections on LASCO-C2 synoptic maps, in: *Proceedings of the SPIE*, 5901, 13–24, 2005.
- Boursier, Y., Lamy, P., Llebaria, A., and Robelus, S.: The Marseille-Artemis Catalog of LASCO CMEs, ESA SP-617, 119, 2006.
- Burkepile, J. T. and St. Cyr, O. C.: A revised and expanded catalogue of coronal mass ejections observed by the Solar Maximum Mission coronagraph, Tech Note TN-369+STR, Natl. Cent. For Atmos. Res., Boulder, Colo., 1993.
- Gopalswamy, N., Yashiro, S., Kaiser, M. L., Howard, R. A., and Bougeret, J.-L.: Radio Signatures of Coronal Mass Ejection Interaction: Coronal Mass Ejection Cannibalism?, *Astrophys. J.*, 548, L91–L94, 2001.
- Gopalswamy, N., Yashiro, S., Michael, G., Kaiser, M. L., Howard, R. A., Reames, D. V., Leske, R., and von Rosenvinge, T.: Interacting Coronal Mass Ejections and Solar Energetic Particles, *Astrophys. J.*, 572, L103–L107, 2002.
- Gopalswamy, N.: A global picture of CMEs in the inner heliosphere, in: *The Sun and the Heliosphere as an Integrated system*, ASSL series, edited by: Poletto, G. and Suess, S., KLUWER/Boston, Chapter 8, p. 201, 2004.
- Gopalswamy, N.: Coronal Mass Ejections of cycle 23, *J. Astrophys. Astron.*, 27, 243–254, 2006a.
- Gopalswamy, N.: Consequences of Coronal Mass Ejections in the Heliosphere, *Sun and Geosphere* (ISSN: 1819-0839), 1, 5–12, 2006b.
- Gopalswamy, N., Shimojo, M., Lu, W., Yashiro, S., Shibasaki, K., and Howard, R. A.: Prominence Eruptions and Coronal Mass Ejection: A Statistical Study using Microwave Observations, *Astrophys. J.*, 586, 562–578, 2003a.
- Gopalswamy, N., Lara, A., Yashiro, S., and Howard, R. A.: Coronal mass ejections and solar polarity reversal, *Astrophys. J.*, 598, L63–L66, 2003b.

- Gopalswamy, N., Yashiro, S., and Akiyama, S.: Coronal mass ejections and space weather due to extreme events, in: *Solar Influence on the Heliosphere and Earth's Environment: Recent Progress and Prospects*, edited by: Gopalswamy, N. and Battacharyya, A., Quest Publications, Mumbai, 79–84, 2006.
- Gopalswamy, N., Yashiro, S., Michalek, G., Stenborg, G., Vourlidas, A., Freeland, S., and Howard, R. A.: The SOHO/LASCO CME CATALOG, Earth, Moon, and Planets, in press, 2008.
- Howard, R. A., Sheeley, N. R., Michels, D. J., and Koomen, M. J.: Coronal mass ejections – 1979–1981, *J. Geophys. Res.*, 90, 8173–8191, 1985.
- Hundhausen, A. J.: Sizes and Locations of Coronal Mass Ejections: SMM Observations From 1980 and 1984–1989, *J. Geophys. Res.*, 98, 13 117–13 200, 1993.
- Qu, M., Shih, F. Y., Jing, J., and Wang, H.: Automatic Detection and Classification of Coronal Mass Ejections, *Solar Phys.*, 237, 419–431, 2006.
- Robbrecht, E. and Berghmans, D.: Automated recognition of coronal mass ejections (CMEs) in near-real-time data, *Astron. Astrophys.*, 425, 1097–1106, 2004.
- Robbrecht, E., Berghmans, D., and van der Linden, R. A. M.: Objective CME detection over the solar cycle: A first attempt, *Adv. Space Res.*, 38, 475–479, 2006.
- St. Cyr, O. C. and Burkepile, J. T.: A catalogue of mass ejections observed by the solar maximum mission coronagraph, *Tech. Note NCAR/TN-352+STR*, Natl. Cent. for Atmos. Res., Boulder, Colorado, 1990.
- St. Cyr, O. C., Howard, R. A., Sheeley, N. R., Plunkett, S. P., Michels, D. J., Paswaters, S. E., Koomen, M. J., Simnett, G. M., Thompson, B. J., Gurman, J. B., Schwenn, R., Webb, D. F., Hildner, E., and Lamy, P. L.: Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998, *J. Geophys. Res.*, 105, 18 169–18 186, 2000.
- Tousey, R.: The solar corona, in: *Space Research XIII*, edited by: Rycroft, M. J. and Runcorn, S. K., 713–730, 1973.
- Yashiro, S., Gopalswamy, N., Michalek, G., St. Cyr, O. C., Plunkett, S. P., Rich, N. B., and Howard, R. A.: A catalog of white light coronal mass ejections observed by the SOHO spacecraft, *J. Geophys. Res.*, 109, A07105, doi:10.1029/2003JA010282, 2004.
- Yashiro, S., Gopalswamy, N., Akiyama, S., Michalek, G., and Howard, R. A.: Visibility of coronal mass ejections as a function of flare location and intensity, *J. Geophys. Res.*, 110, A12S05, doi:10.1029/2005JA011151, 2005.
- Yashiro, S., Akiyama, S., Gopalswamy, N., and Howard, R. A.: Different power-law indices in the frequency distributions of flares with and without coronal mass ejections, *Astrophys. J.*, 650, L143–L146, 2006.
- Yurchyshyn, V., Yashiro, S., Abramenko, V., Wang, H., and Gopalswamy, N.: Statistical Distributions of Speeds of Coronal Mass Ejections, *Astrophys. J.*, 619, 599–603, 2005.